Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS Advance Access published September 28, 2013
MNRAS (2013) doi:10.1093/mnras/stt1643

Cosmological constraints from applying SHAM to rescaled cosmological simulations

Vimal Simha⋆ and Shaun Cole
Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

Accepted 2013 August 28. Received 2013 August 27; in original form 2013 February 7

ABSTRACT

We place constraints on the matter density of the Universe and the amplitude of clustering using measurements of the galaxy two-point correlation function from the Sloan Digital Sky Survey (SDSS). We generate model predictions for different cosmologies by populating rescaled *N*-body simulations with galaxies using the subhalo abundance matching (SHAM) technique. We find $\Omega_{\rm M} = 0.29 \pm 0.03$ and $\sigma_8 = 0.86 \pm 0.04$ at 68 per cent confidence by fitting the observed two-point galaxy correlation function of galaxies brighter than $M_r = -18$ in a volume-limited sample of galaxies obtained by the SDSS. We discuss and quantify potential sources of systematic error and conclude that while there is scope for improving its robustness, the technique presented in this paper provides a powerful low-redshift constraint on the cosmological parameters that is complementary to other commonly used methods.

Key words: methods: numerical – galaxies: haloes – cosmological parameters – large-scale structure of Universe.

1 INTRODUCTION

Determining cosmological parameters such as the matter density of the Universe and the amplitude of clustering is a key goal of cosmology. Measurements of the clustering of galaxies provide a unique window into the distribution of dark matter in the Universe from which cosmological parameters can be inferred. Galaxy surveys map the distribution of visible baryons which indirectly trace the underlying distribution of the gravitationally dominant dark matter. While the distribution of dark matter for a given cosmology can be reliably computed from first principles using cosmological *N*-body simulations, because of uncertainties in the detailed physics of galaxy formation – gas cooling, star formation, feedback, etc. – the distribution of galaxies cannot be robustly predicted from first principles. Consequently, one of the principal impediments in inferring cosmological information from observations of galaxy clustering is galaxy bias, the difference between the distribution of galaxies and the underlying dark matter. In this paper, we use a simple non-parametric model to relate observed galaxy luminosity to halo mass in an *N*-body simulation and use this model to place constraints on the universal matter density ($\Omega_{\rm M}$) and the amplitude of clustering (σ_8) by fitting the observed clustering of galaxies.

We relate (sub)haloes in our *N*-body simulation to observed galaxies using subhalo abundance matching (SHAM). SHAM is a simple, non-parametric model based on assuming a monotonic relationship between observed galaxy luminosity and simulated (sub)halo mass. In this model, all galaxies are assumed to be contained within dark matter subhaloes, and galaxy luminosity is assumed to be monotonically related to the present-day subhalo mass for central subhaloes and the subhalo mass at the accretion epoch for satellite subhaloes. SHAM takes only the space density of galaxies as input and predicts the clustering of a galaxy population. We use it to predict the clustering of observed volume-limited samples of galaxies. Distinct cosmologies produce distinct populations of dark matter haloes and subhaloes (Zheng et al. 2002); thus, for each cosmology, the SHAM model predicts a distinct galaxy autocorrelation function.

SHAM has been successfully used to match theoretical predictions to observables. For example, Guo et al. (2010) match the observed stellar mass–halo mass relation by populating *N*-body simulations using the SHAM technique. Trujillo-Gomez et al. (2011) match the observed circular velocity statistics using a similar method. Simha et al. (2012) found good agreement between its assumptions and the output from a cosmological smoothed particle hydrodynamics (SPH) simulation which incorporated gas physics, star formation and feedback from supernovae-driven winds. While SHAM has mainly been used for galaxy formation applications, Vale & Ostriker (2006) employ the technique to draw inferences about cosmological parameters finding that the *Wilkinson Microwave Anisotropy Probe* 1 (*WMAP*1) cosmology with $\Omega_{\rm M} = 0.3$ and $\sigma_8 = 0.9$ predicts galaxy cluster mass-to-light ratios that are too high compared to the observations.

Running a suite of high-resolution cosmological *N*-body simulations to adequately sample the cosmological parameter space is computationally demanding. Instead, we rescale the masses, positions and velocities of haloes/subhaloes obtained from one

⋆E-mail: vimal.simha@durham.ac.uk

simulation to different target cosmologies. We use the reduced version of the rescaling technique of Angulo & White (2010) as implemented by Ruiz et al. (2011) to construct subhalo catalogues for a given set of cosmological parameters. Several authors (Angulo & White 2010; Ruiz et al. 2011; Guo et al. 2013) have shown that the relevant properties of (sub)haloes in such scaled models are very close to those in simulations carried out directly with the target cosmology.

Our approach is similar in spirit to previous studies such as Tinker et al. (2012), who use halo occupation distribution (HOD) models; Cacciato et al. (2013), who use conditional luminosity function (CLF) combined with the halo model; and Harker, Cole & Jenkins (2007), who use semi-analytic models and a similar simulation rescaling technique to fit the observed clustering of galaxies to infer cosmological parameters. In contrast to the semi-analytic models of galaxy formation used by Harker et al. (2007), we do not make any assumptions about the detailed gas physics that is involved in the process of galaxy formation. Although HOD models such as those used by Tinker et al. (2012) are also based on statistical descriptions rather than detailed modelling of the gas physics, there are significant differences between our techniques. HOD models describe galaxy bias using a probability distribution $P(N|M)$, the probability that a halo of mass M contains N galaxies of a given type. In HOD models, the satellite occupation function, the mean number of satellites as a function of halo mass, is constructed using a parametric model. In contrast, in SHAM, the satellite occupation function is determined by the N-body simulation and directly tied to the cosmological model. Because SHAM introduces a stronger prior on the relation between the galaxy and halo populations, it has the potential to get useful cosmological constraints from weaker data.

In Section 2, we describe our methods – data, simulation, SHAM and rescaling technique. In Section 3, we describe our theoretical model and how it is sensitive to different cosmologies. In Section 4, we discuss our results and constraints on cosmological parameters. In Section 5, we discuss systematic uncertainties in our technique. In Section 6, we discuss the utility of our model in fitting other data sets to obtain complementary constraints on the cosmology. Finally, in Section 7, we summarize our results.

2 METHODS

2.1 Data

Our galaxy clustering measurements are obtained from 7900 deg^2 of sky observed by the Sloan Digital Sky Survey (SDSS; Zehavi et al. 2011). We restrict our analysis to volume-limited samples of galaxies.

The clustering quantity we use for each sample is the projected autocorrelation function $w_{\rm p}(r_{\rm p})$ defined as

$$w_{\rm p}(r_{\rm p}) = 2\int_0^{\pi_{\rm max}} \xi(r_{\rm p},\pi)\,{\rm d}\pi, \quad (1)$$

where $r_{\rm p}$ is the projected separation between two galaxies, π is the line-of-sight separation between two galaxies and $\xi(r_{\rm p},\pi)$ is the measured two-dimensional correlation function. Due to the finite volume of the survey, the integral is limited to $\pi_{\rm max} = 40\,h^{-1}$ Mpc.

Our analysis is primarily focused on a volume-limited sample of galaxies brighter than $M_r = -18.0$. In addition, we make use of other volume-limited samples to test the robustness of our conclusions. The error estimates for each clustering sample are computed using the jackknife technique (see Zehavi et al. 2011 for more details).

2.2 Simulations

We use two simulations, the Millennium Simulation (MS; Springel et al. 2005) and the higher resolution Millennium II simulation (MS-II; Boylan-Kolchin et al. 2009). Both simulations follow the evolution of 2160^3 particles from $z = 127$ to 0 in a Λ cold dark matter (ΛCDM) cosmology (inflationary, cold dark matter with a cosmological constant) with $\Omega_{\rm M} = 0.25$, $\Omega_\Lambda = 0.75$, $h \equiv H_0/100$ km s^{-1} Mpc^{-1} $= 0.73$, primordial spectral index $n_{\rm s} = 1$ and the amplitude of mass fluctuations, $\sigma_8 = 0.9$, where σ_8 is the linear theory rms mass fluctuation amplitude in spheres of radius $8\,h^{-1}$ Mpc at $z = 0$. These parameter values were chosen to agree with *WMAP*1 data (Spergel et al. 2003), and are different from but reasonably close to current estimates from the cosmic microwave background (CMB; Larson et al. 2011) and large-scale structure (Reid et al. 2010), the main difference is that the more recent data favour a lower value of σ_8.

The MS simulates a comoving box that is $500\,h^{-1}$ Mpc on each side, while the MS-II simulates a comoving box that is $100\,h^{-1}$ Mpc on each side. The simulation particle masses are $m_{\rm p} = 8.6 \times 10^8\,h^{-1}\,{\rm M}_\odot$ in MS and $6.9 \times 10^6\,h^{-1}\,{\rm M}_\odot$ in MS-II.

For each output epoch in each simulation, the friends-of-friends (FOF) algorithm is used to identify groups by linking together particles separated by less than 0.2 of the mean interparticle separation (Davis et al. 1985). The SUBFIND algorithm (Springel, Yoshida & White 2001) is then applied to each FOF group to split it into a set of self-bound subhaloes. The central subhalo is defined as the most massive subunit of an FOF group. We construct subhalo merger trees which link each subhalo at each epoch to a unique descendent in the following epoch. These merger trees allow us to track the formation history of each (sub)halo that is identified at $z = 0$. Springel et al. (2001) and Boylan-Kolchin et al. (2009) provide a detailed description of these simulations and the post-processing techniques.

2.3 Subhalo abundance matching

SHAM is a technique for assigning galaxies to simulated dark matter haloes and subhaloes. The essential assumptions are that all galaxies reside in identifiable dark matter substructures and that luminosity or stellar mass of a galaxy is monotonically related to the potential well depth of its host halo or subhalo. Some implementations use the maximum of the circular velocity profile as the indicator of potential well depth, while others use halo or subhalo mass. The first clear formulations of SHAM as a systematic method appear in Conroy, Wechsler & Kravtsov (2006) and Vale & Ostriker (2006), but these build on a number of previous studies that either test the underpinnings of SHAM or implicitly assume SHAM-like galaxy assignment (e.g. Colín et al. 1999; Kravtsov et al. 2004; Nagai & Kravtsov 2005).

N-body simulations produce subhaloes that are located within the virial radius of haloes. The present mass of subhaloes is a product of mass built up during the period when the halo evolves in isolation and tidal mass-loss after it enters the virial radius of a more massive halo (e.g. Kazantzidis et al. 2004; Kravtsov et al. 2004). The stellar component, however, is at the bottom of the potential well and more tightly bound making it less likely to be affected by tidal forces. Therefore, several authors (e.g. Conroy et al. 2006; Vale & Ostriker 2006) argue that the properties of the stellar component should be more strongly correlated with the subhalo mass at the epoch of accretion rather than at $z = 0$.

Vale & Ostriker (2006) apply a global statistical correction to subhalo masses relative to halo masses (as do Weinberg et al. 2008), while Conroy et al. (2006) explicitly identify subhaloes at the epoch of accretion and use the maximum circular velocity at that epoch. Our formulation here is similar to that of Conroy et al. (2006), though we use mass rather than circular velocity. Specifically, we assume a monotonic relationship between galaxy luminosity and halo mass at the epoch of accretion and determine the form of this relation by solving the implicit equation

$$n_{\rm S}(< M_r) = n_{\rm H}(> M_{\rm H}), \tag{2}$$

where $n_{\rm S}$ and $n_{\rm H}$ are the number densities of galaxies and haloes, respectively, M_r is the galaxy r-band magnitude threshold and $M_{\rm H}$ is the halo mass threshold chosen so that the number density of haloes above it is equal to the number density of galaxies in the sample. The quantity $M_{\rm H}$ is defined as follows:

$$M_{\rm H} = \begin{cases} M_{\rm halo}(z=0) & \text{for distinct haloes,} \\ M_{\rm halo}(z=z_{\rm sat}) & \text{for subhaloes,} \end{cases} \tag{3}$$

where $z_{\rm sat}$ is the epoch when a halo first enters the virial radius of a more massive halo.

Guo et al. (2013) compare the mass functions of SHAM-selected subhaloes (see their fig. 1) in MS and MS-II finding good convergence over the mass range where both simulations have adequate resolution. At $z = 0$, they find that the (SHAM-selected) subhalo mass function converges for subhaloes with more than $\sim 10^3$ particles at infall. This is more than an order of magnitude greater than the level at which the halo mass function converges (Boylan-Kolchin et al. 2009), because subhaloes experience tidal stripping after infall and must therefore have a larger number of particles at infall to be reliably identified at $z = 0$.

Fig. 1 shows the satellite occupation function as a function of halo mass, i.e. the mean number of SHAM-selected satellite subhaloes per parent halo as a function of parent halo mass. In both MS and MS-II, we rank order subhaloes identified at $z = 0$ by $M_{\rm H}$, mass at $z = 0$ for central subhaloes and at infall for satellite subhaloes. We then select subhaloes above a mass threshold of $8.6 \times 10^{10}\, h^{-1}\, {\rm M}_\odot$ which is 100 times the particle mass in MS. Because of its higher resolution (factor of ~ 100 difference in particle mass), more satellite subhaloes are identified in MS-II compared to MS across the range of parent halo masses.

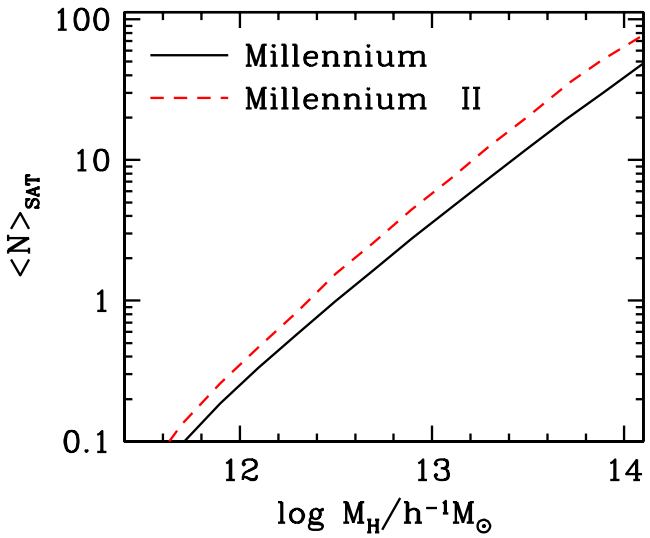

Figure 1. Mean number of SHAM-selected subhaloes per halo as a function of parent halo mass at $z = 0$. Subhaloes above an infall mass threshold of $8.6 \times 10^{10}\, h^{-1}\, {\rm M}_\odot$ are selected which corresponds to 100 times the particle mass in the MS.

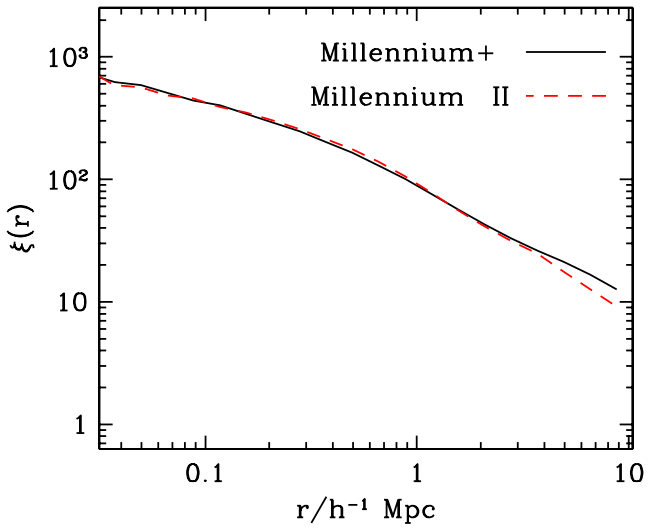

Figure 2. Two-point correlation function of $z = 0$ subhaloes that are above an infall mass threshold. In the MS-II, we only include subhaloes identified at $z = 0$, while in the MS, we add a fraction of disrupted subhaloes to those identified at $z = 0$ so that we match the mean number of subhaloes per parent halo as a function of parent halo mass in MS-II (see Fig. 1).

In MS, there is a missing population of satellite subhaloes that have fewer than 10^3 particles at infall and cannot be identified at $z = 0$ because of subsequent stripping, although similar objects are identified in the higher resolution MS-II. In order to exploit the larger box size and better statistics in MS, we construct a subhalo catalogue, MS Plus (MS+), that matches the satellite occupation function of MS-II. To achieve this, we augment the subhaloes identified at $z = 0$ in MS with additional subhaloes that are identified at high z, but cannot be identified at $z = 0$. These subhaloes are randomly chosen from among disrupted subhaloes that were above the mass threshold at infall. The number of additional subhaloes in each parent halo mass bin is set so as to obtain a match to the MS-II satellite occupation function. The $z = 0$ positions of each disrupted subhalo are obtained by tracking the position of its most bound particle from $z_{\rm sat}$ to $z = 0$.

The difference in the satellite occupation functions between MS and MS-II depends on the mass threshold that is applied, and for a mass threshold above $\sim 10^{12}\, h^{-1}\, {\rm M}_\odot$, they converge. Therefore, the MS+ catalogue has to be freshly constructed for each mass threshold that is required. While fitting the SDSS data, we carry out the procedure to construct MS+ with a mass threshold which gives the galaxy number density of the observed sample.

Fig. 2 shows the two-point correlation function of SHAM-selected subhaloes in MS-II compared to MS+. On scales not affected by box size effects, the autocorrelation functions extracted from the MS-II simulation and our MS+ method described above show strong agreement demonstrating the efficacy of this procedure.

2.4 Rescaling technique

Using the clustering of galaxies at $z = 0$ to infer cosmological parameters requires modelling the growth of density perturbations in the non-linear regime. In principle, one would need a numerical simulation for each set of cosmological parameters. However,

running such a suite of high-resolution cosmological N-body simulations to adequately sample the cosmological parameter space is computationally expensive.

Instead, we rescale our simulation from its 'native' cosmology to different cosmologies with slightly different parameters, using a technique that involves rescaling the box length, velocity and mass units, and relabelling output times. The first attempt to rescale the simulation output to a different cosmology was carried out by Zheng et al. (2002). Harker et al. (2007) used this rescaling technique in combination with semi-analytic models of galaxy formation to constrain the cosmological parameters. More recently, Angulo & White (2010) presented the technique of rescaling the output of a cosmological simulation from its 'native' cosmology to a cosmology with different parameters as a systematic algorithm.

In this paper, we use the Angulo & White (2010) algorithm to generate (sub)halo catalogues for a given set of cosmological parameters. Following Ruiz et al. (2011), we do not apply the final step of the Angulo & White (2010) algorithm which involves adjusting the amplitudes of large-scale linear modes which would in any case have negligible effects on the clustering statistics considered in this paper.

Angulo & White (2010) and Ruiz et al. (2011) find good agreement between the halo catalogues extracted from simulations scaled to a target cosmology and simulations run with the target cosmology. Guo et al. (2013) compare the properties of the subhalo catalogue in a simulation run with the *WMAP*7 cosmology ($\Omega_{\rm M} = 0.27$ and $\sigma_8 = 0.8$) to the subhalo catalogue obtained by rescaling the MS to the *WMAP*7 cosmology using a technique identical to that implemented in this paper. They find very good agreement between the mass functions (Fig. 1) and two-point correlation functions (Fig. 4) of SHAM-selected subhaloes indicating that the rescaling technique implemented in this paper does not introduce significant biases in subhalo positions and masses.

3 COSMOLOGICAL CONSTRAINTS FROM THE GALAXY TWO-POINT CORRELATION FUNCTION

We probe the cosmology by fitting the observed galaxy two-point correlation function of a volume-limited sample of galaxies. To generate a theoretical prediction for the two-point correlation function for a given cosmology, we implement the following steps. First, we rescale our simulation to the target cosmology using the procedure described in Section 2.4. We then select (sub)haloes using the SHAM procedure described in Section 2.3. Briefly, subhaloes above a mass threshold at infall for satellite subhaloes and at $z = 0$ for central subhaloes are selected. Our subhalo mass threshold is set so that the number density of SHAM-selected subhaloes in our simulation is equal to the number density of observed galaxies in the sample we are fitting to (see equation 2). Finally, we compute the projected two-point correlation function of these SHAM-selected subhaloes in the simulation cube.

Fig. 3 provides a pedagogical demonstration of our technique demonstrating that, in principle, it could be used to obtain cosmological constraints. Panel (a) shows the effect of changing $\Omega_{\rm M}$ on the two-point galaxy correlation function in a standard ΛCDM cosmology with all other cosmological parameters held constant. Decreasing $\Omega_{\rm M}$ boosts the two-point correlation function on all scales. Panel (b) shows the effect of changing σ_8 on the two-point galaxy correlation function in a standard ΛCDM cosmology with all other cosmological parameters held constant. Increasing σ_8 boosts the two-point correlation function on all scales.

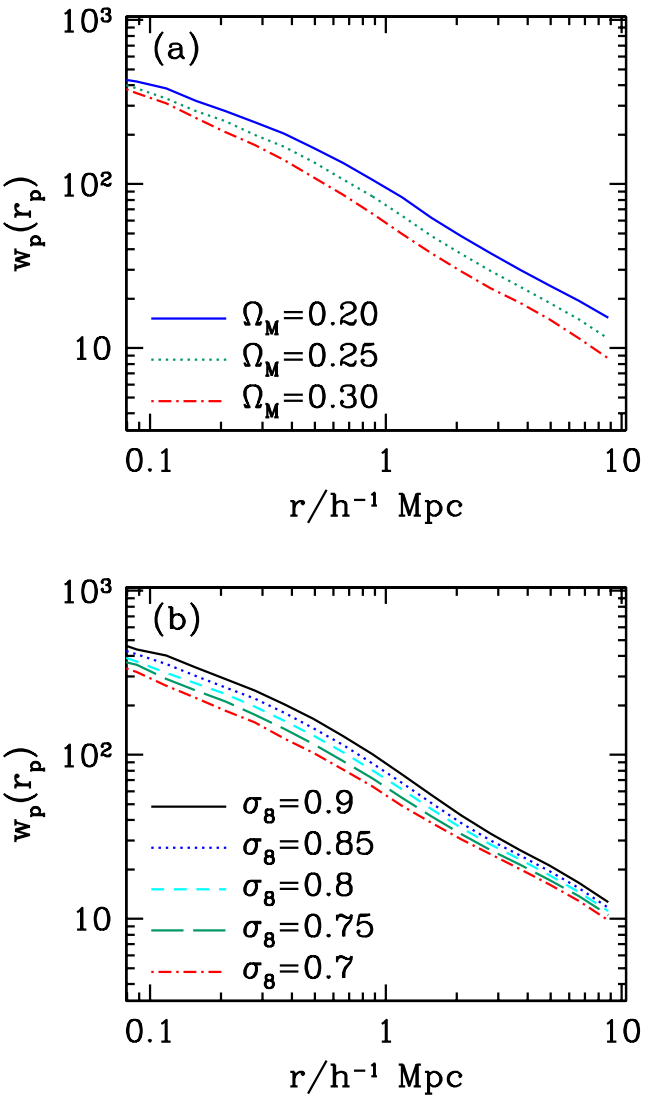

Figure 3. A pedagogical demonstration of the method employed in this paper. The effect on the two-point correlation function of SHAM-selected subhaloes of changing $\Omega_{\rm M}$ at fixed σ_8 in panel (a) and of changing σ_8 at fixed $\Omega_{\rm M}$ in panel (b).

4 COSMOLOGICAL CONSTRAINTS

We determine constraints on the cosmological parameters by comparing the predicted two-point galaxy correlation function for a grid of models in the $\sigma_8 - \Omega_{\rm M}$ plane to the SDSS-observed two-point galaxy correlation function of galaxies brighter than $M_r = -18$. We restrict our analysis to scales below $10\,h^{-1}$ Mpc. We calculate the expected correlation function at $z = 0.1$ as a function of cosmology which corresponds to the median redshift of the data we compare to.

We compute the χ^2 for each model using the full covariance matrix. The covariance matrix is calculated by Zehavi et al. (2011) using jackknife resampling (see section 2.2 of Zehavi et al. 2011 for further details of the method). Besides the two cosmological parameters, $\Omega_{\rm M}$ and σ_8, there are no free parameters in our model.

We impose the following flat priors on our cosmological parameters such that $0.2 \leq \Omega_{\rm M} \leq 0.35$ and $0.65 \leq \sigma_8 \leq 1$. Rescaling our simulation to a model with a different amplitude of clustering relies on relabelling the simulation output epochs. Since our

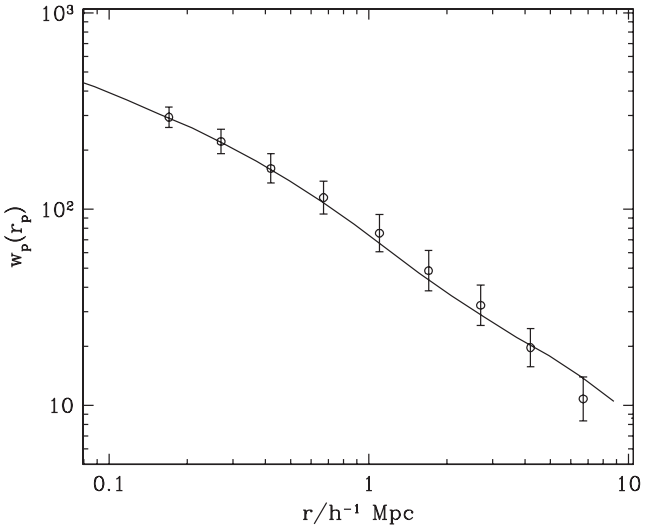

Figure 4. The solid curve shows the galaxy two-point correlation function of our best-fitting model with $\Omega_{\rm M} = 0.275$ and $\sigma_8 = 0.86$. The points with error bars show the SDSS-observed galaxy two-point correlation function from a volume-limited sample of galaxies with $M_r \leq -18.0$.

simulations are only run to $z = 0$, we are restricted to $\sigma_8 \leq 0.9$. For $0.6 \leq \sigma_8 \leq 0.9$, linearly interpolating $\log(\xi)$ with σ_8 works accurately, and so we extend this scaling to extrapolate the galaxy two-point correlation function to generate predictions for models with $\sigma_8 \geq 0.9$.

Fig. 4 shows the galaxy two-point correlation function for our best-fitting model with $\Omega_{\rm M} = 0.275$ and $\sigma_8 = 0.86$ plotted against the data. The χ^2 for our best-fitting model is 9.21. As we have two free parameters and nine data points, this corresponds to a χ^2 per degree of freedom of 1.3. The probability of drawing a value ≥ 9.21 from a χ^2 distribution with 7 degrees of freedom is 0.24, and so our best-fitting model is an acceptable fit to the data.

Fig. 5 shows our main result, cosmological constraints in the $\Omega_{\rm M}$–σ_8 plane. Because of the way $\Omega_{\rm M}$ and σ_8 affect the two-point galaxy correlation function (see Section 3), our constraints on them are correlated. Note that only the portion of the outer contour with

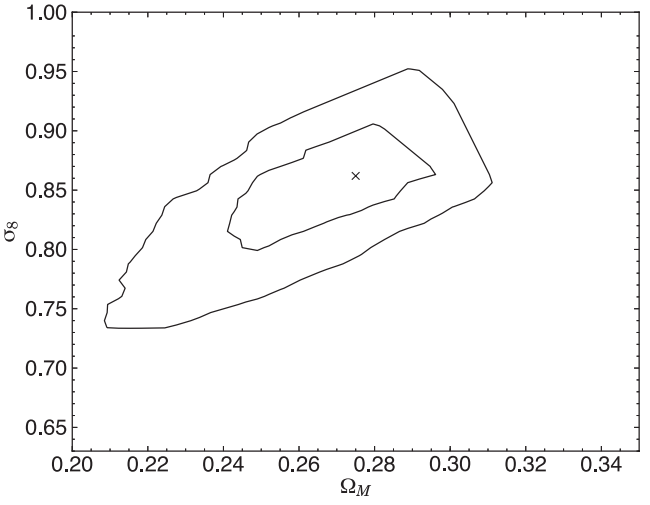

Figure 5. Joint constraint in the σ_8–$\Omega_{\rm M}$ plane. The inner contour shows the boundary of the 68 per cent confidence region and the outer contour shows the 95 per cent confidence region.

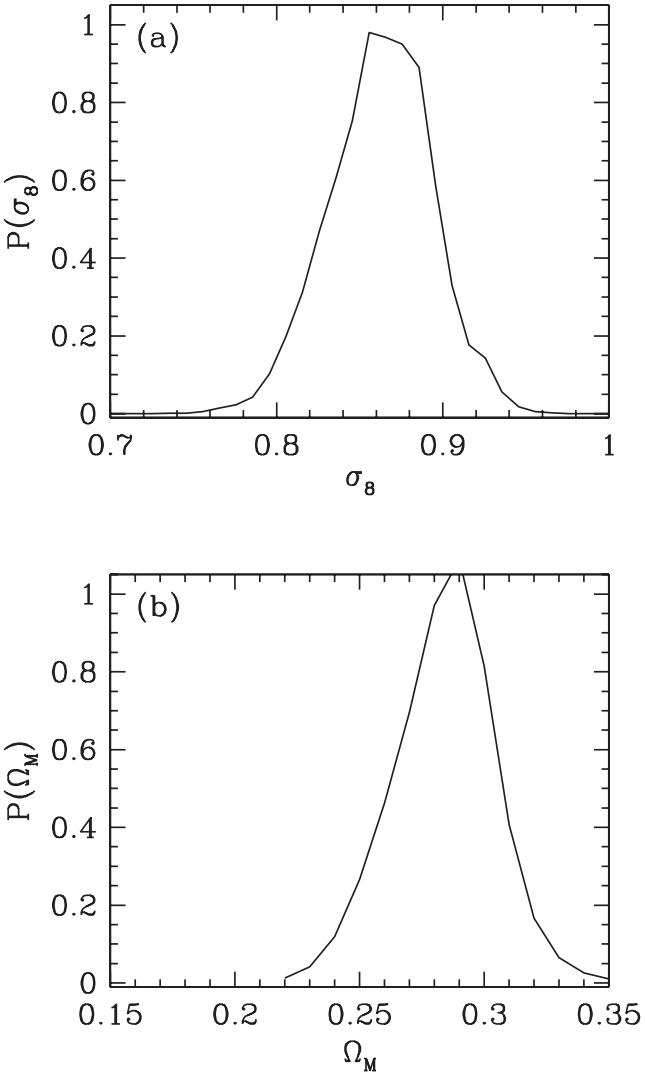

Figure 6. Marginalized posterior probability distribution of σ_8 in panel (a) and of $\Omega_{\rm M}$ in panel (b) from fitting the two-point galaxy correlation function.

$\sigma_8 \geq 0.9$ depends on our extrapolation of the two-point correlation function.

Fig. 6 shows the marginalized probability distributions of $\Omega_{\rm M}$ and σ_8. Our marginalized constraints on the individual parameters are $\Omega_{\rm M} = 0.29 \pm 0.03$ and $\sigma_8 = 0.86 \pm 0.04$ (68 per cent).

We emphasize that our constraints on these parameters are obtained for fixed values of other cosmological parameters. We defer discussion of the implications of this to the final section.

5 SYSTEMATIC UNCERTAINTIES

We discuss the systematic uncertainties in our analysis, focusing on the uncertainties in our theoretical modelling. See Zehavi et al. (2011) for a discussion of observational uncertainties.

5.1 Scatter in SHAM

SHAM assumes a strictly monotonic relation between (sub)halo mass and galaxy r-band luminosity with zero scatter. While the assumption of zero scatter is idealized, several observational studies (e.g. van den Bosch et al. 2007; Zheng, Coil & Zehavi 2007)

indicate that the scatter in luminosity at fixed (sub)halo mass is small. While from the theory side, Simha et al. (2012) find a strong correlation between r-band luminosity and subhalo mass in their SPH simulations, with a small scatter of 0.15 dex in luminosity at fixed subhalo mass.

Scatter in the subhalo mass–luminosity relation affects the correlation function by altering the subhaloes that are included in the sample. The effect of scatter is felt only at the boundary of the sample since assigning galaxies to different haloes/subhaloes within the sample does not change the correlation function. Therefore, the effect of scatter in the luminosity–subhalo mass relation depends on how steep the luminosity–bias relation is at the boundary of the sample. For our sample, random scatter in the subhalo mass–luminosity relation has negligible impact on the galaxy/subhalo two-point correlation function. However, a similar level of scatter in the subhalo mass–luminosity relation will have a larger effect on the correlation function at a brighter luminosity threshold where the luminosity–bias relation is steeper. Additionally, scatter could potentially be a significant source of uncertainty for other statistics that are more sensitive to accurate identification of the host haloes of galaxies.

5.2 Satellite galaxy fraction

SHAM relies on accurately resolving and identifying substructures and recovering their properties at infall. Since we trace the $z_{\rm sat}$ progenitors of $z = 0$ substructures, our results are unlikely to be affected by random fluctuations in the density field of haloes that may be spuriously identified as substructures. However, if subhaloes hosting satellite galaxies that have merged with the central galaxy of the halo are identified as substructures, we would overestimate the halo occupation of massive haloes. Conversely, subhaloes that fall into more massive haloes and lose a substantial fraction of their mass due to tidal stripping may no longer be resolved in the simulation at $z = 0$, although they may still host satellite galaxies.

Simha et al. (2012) find that for a fixed subhalo mass, satellite galaxies in their SPH simulation are typically somewhat less luminous than central galaxies because of differences in the ages of their respective stellar populations. Compared to central galaxies of similar stellar mass, satellite galaxies have an older stellar population and are therefore less luminous. If this were to be true of the real Universe, SHAM would underestimate the luminosities of central galaxies and overestimate the luminosities of satellite galaxies, resulting in more satellite galaxies being included in a sample above a given luminosity threshold and consequently an overestimate of the satellite galaxy fraction.

Fig. 7 shows the fraction of galaxies that are satellites, $f_{\rm sat} = n_{\rm sat}/n_{\rm gal}$, where $n_{\rm sat}$ is the number of satellite galaxies and $n_{\rm gal}$ is the total number of galaxies, in bins of (sub)halo mass at infall. Our best-fitting model has an overall $f_{\rm sat} = 0.3$. This is in good agreement with Zehavi et al. (2011), who find $f_{\rm sat} = 0.32 \pm 0.02$ by fitting HOD models to the correlation function of the same sample of galaxies used in this paper. However, the satellite fraction found by Zehavi et al. (2011) is at fixed cosmology, and as we discuss below, there is a degree of degeneracy between the satellite fraction and the cosmological parameters.

Fig. 8 shows the effect of changing $f_{\rm sat}$ on the galaxy correlation function. To increase $f_{\rm sat}$, we include additional 'merged' subhaloes, i.e. subhaloes that exist at high redshift, but have merged by $z = 0$ using its most bound particle when it was last identified to track its position to $z = 0$. We then compute the two-point correlation function for this sample of subhaloes. To decrease $f_{\rm sat}$, we eliminate a fraction of subhaloes chosen randomly.

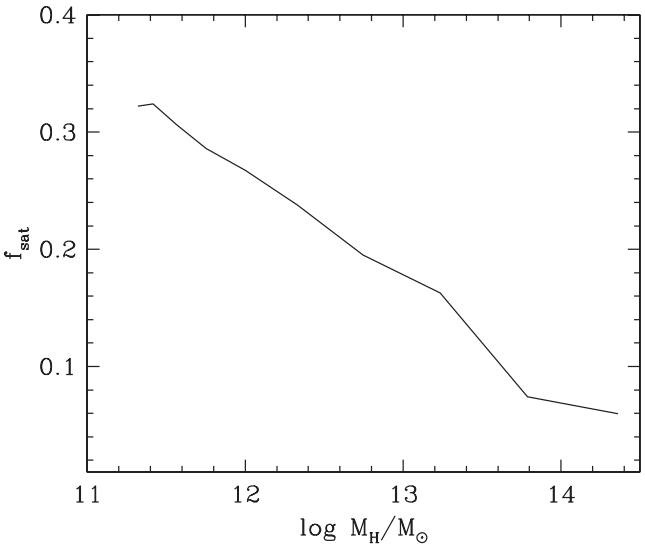

Figure 7. Fraction of galaxies that are satellites, $f_{\rm sat} = n_{\rm sat}/n_{\rm gal}$, where $n_{\rm sat}$ is the number of satellite galaxies and $n_{\rm gal}$ is the total number of galaxies, in bins of (sub)halo mass at infall as a function of (sub)halo mass at infall.

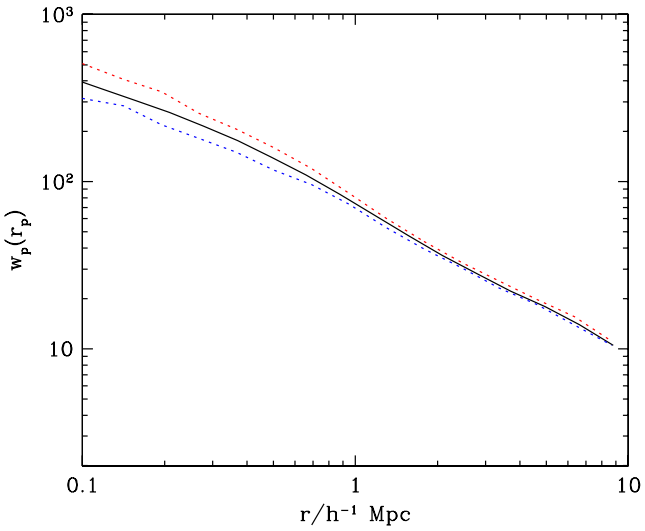

Figure 8. Effect of changing the fraction of satellite galaxies on the galaxy two-point correlation function. The black solid curve shows our fiducial model, while the red and blue dotted curves show the models with a 10 per cent higher and lower $f_{\rm sat}$ compared to the fiducial model, respectively ($f_{\rm sat} = 0.30 \pm 0.03$).

The galaxy correlation function on scales below $\sim 1\,h^{-1}$ Mpc is determined by pairs of galaxies within the same host halo. A higher $f_{\rm sat}$ increases the number of close pairs, boosting the two-point correlation function on small scales. On larger scales, the galaxy correlation function is determined by pairs of galaxies in different haloes. Although the effect of $f_{\rm sat}$ on these scales is much smaller, the correlation function is higher in models with higher $f_{\rm sat}$ because highly biased objects in massive haloes are present in a larger number of pairs compared to a model with lower $f_{\rm sat}$.

The effect on the galaxy correlation function of changing the cosmological parameters shows some degeneracy with the effect of changing $f_{\rm sat}$. At fixed $\Omega_{\rm M}$, a 10 per cent increase in $f_{\rm sat}$ from our fiducial value of $f_{\rm sat} = 0.3$ to 0.33 changes the best-fitting value of σ_8 from 0.86 to 0.83.

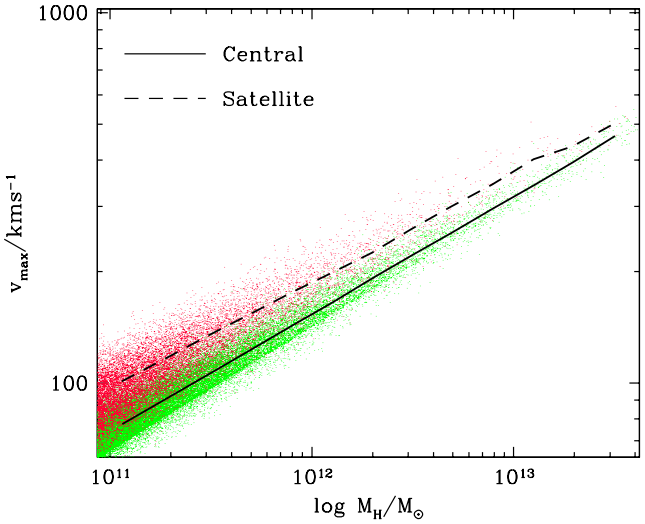

Figure 9. The maximum of the circular velocity profile, $v_{\rm max}$, versus the halo mass, $M_{\rm H}$; both measured at infall for subhaloes and at $z = 0$ for independent haloes. Each green point represents a central galaxy and each red point represents a satellite galaxy. The solid and dashed curves show the mean $v_{\rm max}$ in bins of $M_{\rm H}$ for central and satellite galaxies, respectively.

Our analysis so far has assumed that galaxy luminosity is monotonically related to subhalo mass at infall. Using a different subhalo property as a luminosity proxy would change $f_{\rm sat}$ and hence alter the predicted galaxy correlation function for a given set of cosmological parameters.

Reddick et al. (2012) find large differences in the predicted galaxy two-point correlation function between SHAM-like models that use subhalo properties at infall and models that use $z = 0$ subhalo properties. Because subhaloes undergo substantial tidal stripping after infall, while the more centrally concentrated galaxies at their centres do not, M_0 ($z = 0$ mass) and v_0 ($z = 0$ circular velocity) would both underestimate satellite galaxy luminosity and hence $f_{\rm sat}$ severely, making them unsuitable as galaxy luminosity proxies.

A more realistic alternative galaxy luminosity proxy to subhalo mass at infall ($M_{\rm H}$) is $v_{\rm max}$, the maximum of the circular velocity profile at infall for subhaloes and at $z = 0$ for independent haloes. Fig. 9 shows $v_{\rm max}$ at infall for satellites and at $z = 0$ for central galaxies plotted against $M_{\rm H}$. Each point in the figure represents an individual galaxy, while the solid and dashed curves show the mean $v_{\rm max}$ in bins of $M_{\rm H}$ for central and satellite galaxies. Generally, $v_{\rm max}$ is well correlated with $M_{\rm H}$ in our simulation for both satellite and central galaxies considered independently. However, satellite subhaloes have systematically higher $v_{\rm max}$ than central haloes at fixed $M_{\rm H}$ across the range of halo masses in our simulation. This could be because less concentrated subhaloes are more easily disrupted and the surviving subhaloes have higher than average concentrations. Alternatively, it could be because subhaloes formed earlier than central haloes of similar mass and are hence more concentrated.

Using $v_{\rm max}$ instead of $M_{\rm H}$ for abundance matching at fixed cosmology (fixed at our best-fitting model for $M_{\rm H}$ abundance matching) increases the satellite fraction $f_{\rm sat}$ from 0.3 to 0.38, which is 3σ away from the value of $f_{\rm sat}$ found by Zehavi et al. (2011) by fitting HOD models to the two-point correlation function of the same sample of galaxies.

When abundance matching using $v_{\rm max}$ as the galaxy luminosity proxy, the effect of the large increase in $f_{\rm sat}$ on the galaxy correlation function on large scales can be offset by changing the cosmological parameters. However, on scales below 0.4 h^{-1} Mpc, abundance matching using $v_{\rm max}$ significantly overpredicts the galaxy correlation function, and we are unable to obtain a match to the data for the range of cosmological parameters we consider.

We also consider a model where we use $v_{\rm max}$ as a luminosity proxy for abundance matching, but enforce $f_{\rm sat} = 0.3$ (value of $f_{\rm sat}$ obtained from abundance matching using $M_{\rm H}$ as a luminosity proxy). We create a sample with $f_{\rm sat} = 0.3$ by eliminating satellite subhaloes below a $v_{\rm max}$ threshold at $z = 0$ and including additional central subhaloes in their place to preserve the overall number density of subhaloes. This yields a very similar galaxy autocorrelation function to that obtained from abundance matching using $M_{\rm H}$ as the galaxy luminosity proxy. For $f_{\rm sat} = 0.3$, abundance matching using $v_{\rm max}$ as the luminosity proxy yields cosmological constraints ($\Omega_{\rm M} = 0.28$ and $\sigma_8 = 0.84$) that are consistent with our results obtained with abundance matching using $M_{\rm H}$ as the galaxy luminosity proxy to within 0.5σ.

5.3 Rescaling technique

Ruiz et al. (2011) compare the properties of haloes in a rescaled simulation to a simulation run with the target cosmological parameters. They find that over 99 per cent of haloes with more than 50 particles are recovered in the rescaled simulation. In the rescaled simulations, the masses of haloes are systematically underestimated by ∼5 per cent. But because we use only halo mass to assign a rank to subhaloes, we are not affected by this systematic bias. However, it would be a source of error if a statistic that depended directly on halo mass were being used. The rescaled halo mass–'native' halo mass relation displays scatter. For our purpose, it would mimic the effect of scatter in the subhalo mass–luminosity relation discussed earlier (Section 5.1). The positions of haloes in their rescaled simulations are recovered to a precision of within 100 h^{-1} kpc.

5.4 Other galaxy samples

Our cosmological constraints have been obtained by fitting our model to the observed clustering of one volume-limited sample of galaxies with $M_r \leq -18$. As a consistency check, we compare the clustering predictions of our model with four other volume-limited samples of galaxies taken from SDSS (Zehavi et al. 2011) with $M_r \leq -18.5, -19, -19.5$ and -20.5 corresponding to a number density of galaxies of 2.311, 1.676, 1.12 and 0.318 10^{-2} h^{-3} Mpc^3, respectively. The sample with $M_r \leq -20.5$ corresponds approximately to galaxies brighter than L_*, the characteristic galaxy luminosity above which the number density of galaxies falls exponentially.

For each of these samples, we generate the predicted galaxy two-point correlation function for our best-fitting cosmological model by measuring the clustering of SHAM-selected subhaloes above an infall mass threshold determined so that the number density of subhaloes is equal to the number density of galaxies in the sample. Each panel of Fig. 10 shows the galaxy two-point correlation function predicted by our best-fitting model to the $M_r \leq -18$ sample plotted against the observed galaxy two-point correlation function for the corresponding sample. Formally, our model is a good fit to the data for the volume-limited samples with $M_r \leq -18.5, -19$ and -19.5, but not for the brightest sample with $M_r \leq -20.5$.

The error bars in each panel show the diagonal errors on the observed correlation function. For our main sample and samples with relatively high number density, the error on the correlation function extracted from our simulation is negligible compared to

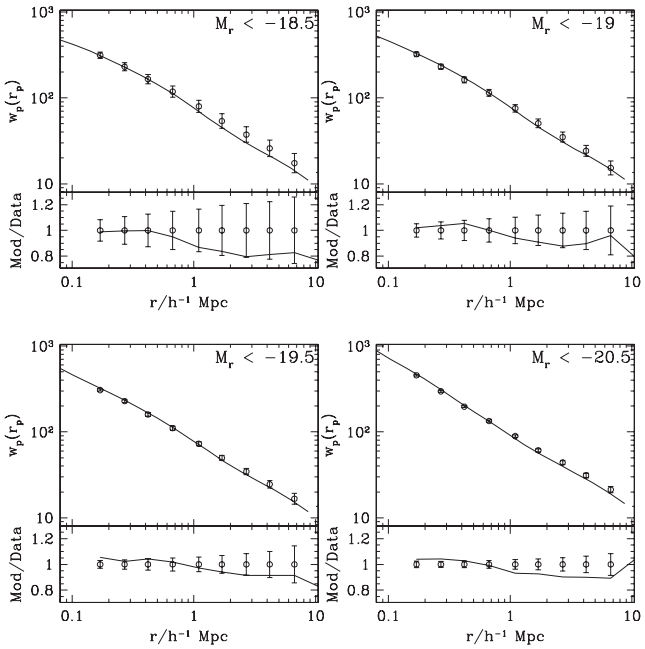

Figure 10. In each panel, the points with error bars show the SDSS-observed galaxy two-point correlation function in a volume-limited sample of galaxies brighter than $M_r = -18.5, -19, -19.5$ and -20.5. The solid curve in each panel shows the galaxy two-point correlation function predicted by our best-fitting model with $\Omega_{\rm M} = 0.275$ and $\sigma_8 = 0.86$ for the corresponding galaxy sample. Lower windows of each panel show the ratio of the model correlation function to the data.

the observational errors and can therefore be ignored. However, because of our finite simulation volume, as we go to brighter samples with low number density, the jackknife estimates of the errors on our predicted correlation function are sufficiently large to be comparable to the observational errors.

When we fit models to these other galaxy samples by allowing the cosmological parameters to vary, we find that for each sample, the best-fitting cosmological parameters are slightly different, but consistent with the constraints from our main sample ($M_r \leq -18$). In principle, tighter constraints could be obtained from a joint fit to all the data sets, but this would require an estimate of their covariance.

6 COMPLEMENTARY CONSTRAINTS ON COSMOLOGY

As we discussed earlier (see Fig. 5), our constraints on $\Omega_{\rm M}$ and σ_8 are correlated because of the way the changes in these parameters affect the galaxy autocorrelation function (see Fig. 3). To a certain extent, the effect of a higher $\Omega_{\rm M}$ can be compensated for by a higher σ_8 if only the galaxy two-point correlation function is used to constrain the cosmology. But complementary constraints on the cosmology can be obtained by probing other observables for which predictions can be generated using the same SHAM model used in this paper.

Fig. 11 shows the effect of the cosmological parameters on the mean number of satellite galaxies per halo, $\langle N \rangle_{\rm sat}$, as a function of halo mass in our simulation using our rescaling technique and the SHAM model. The number of satellite galaxies in a given halo, $N_{\rm sat} = N - 1$, where N is the number of galaxies in the halo. Panel (a) shows the effect of changing $\Omega_{\rm M}$ at fixed σ_8. Changing $\Omega_{\rm M}$ affects the halo mass function making haloes with a given number of satellites more massive for higher $\Omega_{\rm M}$. Panel (b) shows the effect

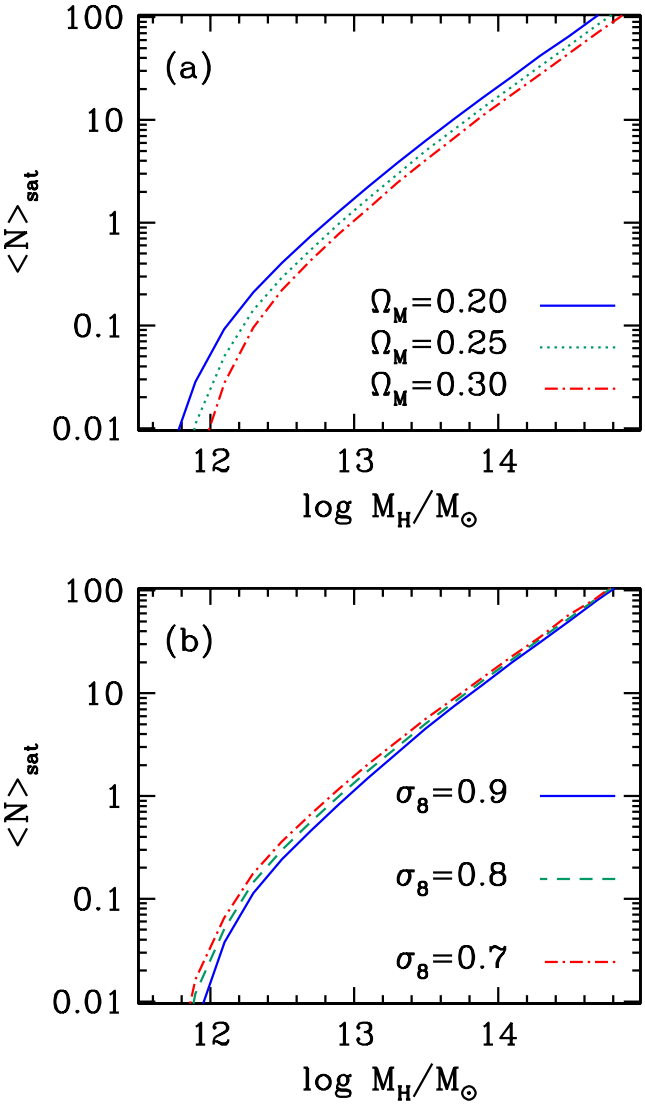

Figure 11. Panel (a) shows the effect of changing $\Omega_{\rm M}$ at fixed σ_8 and panel (b) shows the effect of changing σ_8 at fixed $\Omega_{\rm M}$ on the mean number of satellite galaxies per halo as a function of halo mass.

of changing σ_8 at fixed $\Omega_{\rm M}$. Increasing σ_8 increases the number of haloes of a given mass. Therefore, for a fixed number density of galaxies, there must be fewer galaxies in each halo compared to models with lower σ_8.

Decreasing σ_8 and decreasing $\Omega_{\rm M}$, both boost $\langle N \rangle_{\rm sat}$ rather than counteracting each other as they do for the two-point correlation function. Moving along the degeneracy curve in the $\Omega_{\rm M}$–σ_8 plane in Fig. 5 would generate different and easily distinguishable distributions of $\langle N \rangle_{\rm sat}$. Therefore, for a given cosmology, the additional requirement of matching the galaxy two-point correlation function strongly constrains the HOD, and consequently, a direct measure of the HOD would place constraints on the cosmology that are complementary to our constraints from fitting the two-point correlation function.

Tinker et al. (2012) identify one possible direct measure of the HOD, the mean number of galaxies in haloes of a given mass and use the ratio of these quantities, M/N, to place constraints on cosmological parameters.

7 DISCUSSION AND CONCLUSIONS

We have placed constraints on σ_8 and Ω_M by comparing the SDSS observed projected galaxy two-point correlation function for a volume-limited sample of galaxies with $M_r \leq -18$ to our model predictions generated using *N*-body simulations rescaled to the target cosmology using the technique of Angulo & White (2010) and populated with galaxies using SHAM.

Assuming a flat ΛCDM cosmology with $n_S = 1$, we find $\Omega_M = 0.29 \pm 0.03$ and $\sigma_8 = 0.86 \pm 0.04$ at 68 per cent confidence.

Fig. 12 compares our constraint in the $\Omega_M - \sigma_8$ plane to constraints from *WMAP*7 (Komatsu et al. 2011). Our estimates of both σ_8 and Ω_M are high compared to *WMAP*7, but are consistent at the $\sim 2\sigma$ level.

Our constraints are obtained for fixed values of other cosmological parameters. In contrast, the *WMAP*7 constraints are obtained by marginalizing over all other cosmological parameters. One significant implication of this is in regard to the shape of the power spectrum. We have assumed that the value of the primordial spectral index $n_S = 1$ in our simulation. In contrast, the *WMAP*7 best-fitting value of n_S is 0.96, and $n_S = 1$ is excluded at more than 2σ. While we are unable to comment on the effect of setting $n_S = 0.96$, as this would require new *N*-body simulations with $n_S = 0.96$, on our constraints, forcing $n_S = 1$ while fitting the CMB data would result in a higher best-fitting value of σ_8.

Our results are consistent with and comparable to Tinker et al. (2012), who fit the SDSS galaxy two-point correlation function and M/N (cluster mass to number ratio) using their HOD models, finding $\sigma_8 = 0.85 \pm 0.06$ and $\Omega_M = 0.29 \pm 0.03$.

However, there is some tension between our results and those of Harker et al. (2007), who use semi-analytic models to populate *N*-body simulations rescaled to a given cosmology using a technique similar to that of this paper and fit to the SDSS clustering finding $\sigma_8 = 0.97 \pm 0.06$. In contrast to Harker et al. (2007), who use a semi-analytic model of galaxy formation to populate their *N*-body simulation with galaxies, we use SHAM which only assumes a monotonic relationship between galaxy luminosity and subhalo mass at infall. Secondly, the resolution of MS-II is ∼2000 times higher than the simulations used by Harker et al. (2007). Thirdly, Harker et al. (2007) use a Monte Carlo scheme to generate a merger tree for a halo based on its mass. Consequently, satellite galaxy positions are not obtained from the simulation, and they are instead placed on random particles within the halo. We are unable to quantify the effects of each of these factors. However, repeating the work of Harker et al. (2007) with a high-resolution *N*-body simulation and merger trees and subhalo positions extracted from the *N*-body simulation would be interesting and could potentially reveal the source of the tension.

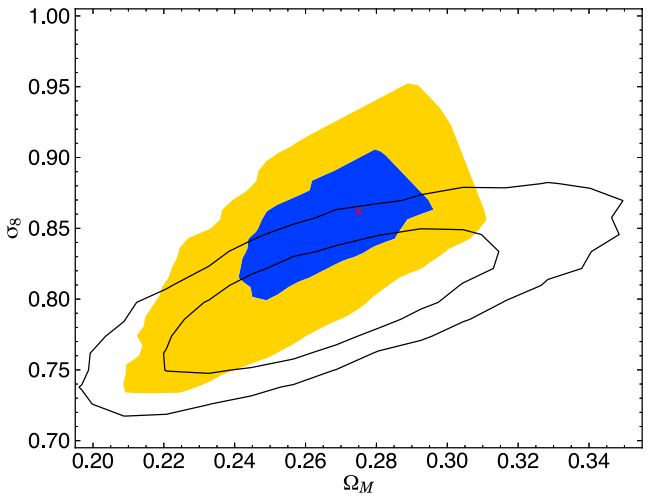

Figure 12. Joint constraint in the $\sigma_8 - \Omega_M$ plane. The inner contour shows the boundary of the 68 per cent confidence region and the outer contour shows the 95 per cent confidence region. The filled contour is the result from this work, while the black solid open contours are from *WMAP*7 (Komatsu et al. 2011).

Besides the CMB and galaxy clustering data, several other methods have been employed to constrain σ_8. For example, Mandelbaum et al. (2012) find $\sigma_8(\Omega_M/0.25)^{0.57} = 0.80 \pm 0.05$ using galaxy–galaxy lensing, and Seljak et al. (2005) find $\sigma_8 = 0.90 \pm 0.03$ by combining their analysis of the Lyman α forest power spectrum with CMB results.

Because our simulation is only run to $z = 0$, we are unable to rescale our subhalo catalogue to cosmologies with $\sigma_8 \geq 0.9$. Our model predictions for $\sigma_8 \geq 0.9$ are based on extrapolating the correlation function as a function of the cosmological parameters. While this appears to be a reasonable approximation, it is likely that it has shortcomings that will be exposed by a simulation that is run with a higher value of σ_8.

We have examined potential sources of systematic uncertainties in our technique, both relating to SHAM and the method we employ for rescaling our simulations to different cosmologies, in Section 5. The most significant source of systematic uncertainty in our technique is the fraction of satellite galaxies, f_{sat}, which could either be underestimated or overestimated in the models, either for numerical reasons relating to resolution and identification of subhaloes or due to systematic differences in the growth of central and satellite galaxies that violate the implicit assumptions of SHAM. The effect of small changes in f_{sat} on the galaxy correlation function can be offset by changes to the cosmological parameters. For example, a 10 per cent increase in f_{sat} from our fiducial value of $f_{sat} = 0.3$ to 0.33 changes the best-fitting value of σ_8 from 0.86 to 0.83 at fixed Ω_M. However, we cannot obtain a good fit to the observed galaxy correlation function if f_{sat} is substantially different from the SHAM prediction. Another source of uncertainty is scatter in the galaxy luminosity–halo mass relation which is assumed to be perfectly monotonic in SHAM. Random scatter in the luminosity–halo mass relation does not affect our cosmological constraints significantly. However, if the scatter were to be correlated with the large-scale environment (beyond the halo), the impact on our results could be significant. The two-point galaxy correlation function on small scales is sensitive to the distribution of galaxies within massive haloes. Therefore, our results would be significantly affected if the distribution of galaxies in haloes were to systematically differ from the distribution of SHAM-selected subhaloes in *N*-body simulations.

Despite these caveats, we emphasize that the technique presented in this paper can provide tight constraints on the cosmology using only low-z data. The remarkable tightness of our constraint arises from the fact that unlike statistical descriptions of the distribution of galaxies in haloes provided by HOD models or the CLF, we do not have the freedom to define the HOD. Furthermore, our model does not make any assumptions about galaxy bias or the detailed physics of galaxy formation except for requiring a monotonic relationship between galaxy luminosity and subhalo mass at infall. Additionally, we assume that galaxies are the observational counterparts of subhaloes identified at $z = 0$ in an *N*-body simulation and that each observed galaxy can be associated with a subhalo.

While future redshift surveys will provide more volume than the SDSS, the potential for tightening of the parameter constraints is limited because of the current systematic uncertainties involved in this technique. However, substantial improvements in the robustness of this technique can be achieved by future investigations. First by adopting the shape of the power spectrum $P(k)$ inferred by the latest CMB observations. Secondly, a large volume simulation with the resolution of MS-II or higher would remove the need for the correction to the subhalo fraction that we apply in this work and the uncertainties associated with it. Systematic uncertainties associated with the rescaling technique scale with the magnitude of the difference in cosmological parameters between the 'native' and rescaled cosmology. To minimize this, it would be useful to carry out a suite of simulations with different cosmological parameters that span the requisite range and only apply the rescaling technique to generate predictions for intermediate values of the parameters. Finally, further investigations of the SHAM technique, particularly with respect to scatter in the luminosity–halo mass relation and the distribution of subhaloes within haloes will help to clarify potential sources of systematic errors involved in this technique.

ACKNOWLEDGEMENTS

We are grateful to John Helly for technical assistance at various stages of this work. We thank Idit Zehavi for providing the covariance matrices. We thank Carlton Baugh for providing invaluable technical assistance, and Carlos Frenk and David Weinberg for useful discussions. The analyses presented in this paper used the Cosmology Machine supercomputer at the ICC, which is part of the DiRAC Facility jointly funded by STFC, the Large Facilities Capital Fund of BIS and Durham University.

REFERENCES

Angulo R. E., White S. D. M., 2010, MNRAS, 405, 143
Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, MNRAS, 398, 1150
Cacciato M., van den Bosch F. C., More S., Mo H., Yang X., 2013, MNRAS, 430, 767
Colín P., Klypin A. A., Kravtsov A. V., Khokhlov A. M., 1999, ApJ, 523, 32
Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111
Guo Q., White S., Angulo R. E., Henriques B., Lemson G., Boylan-Kolchin M., Thomas P., Short C., 2013, MNRAS, 428, 1351
Harker G., Cole S., Jenkins A., 2007, MNRAS, 382, 1503
Kazantzidis S., Mayer L., Mastropietro C., Diemand J., Stadel J., Moore B., 2004, ApJ, 608, 663
Komatsu E. et al., 2011, ApJS, 192, 18
Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35
Larson D. et al., 2011, ApJS, 192, 16
Mandelbaum R., Slosar A., Baldauf T., Seljak U., Hirata C. M., Nakajima R., Reyes R., Smith R. E., 2012, MNRAS, 432, 1544
Nagai D., Kravtsov A. V., 2005, ApJ, 618, 557
Reddick R. M., Wechsler R. H., Tinker J. L., Behroozi P. S., 2012, ApJ, 771, 30
Reid B. A. et al., 2010, MNRAS, 404, 60
Ruiz A. N., Padilla N. D., Domínguez M. J., Cora S. A., 2011, MNRAS, 418, 2422
Seljak U. et al., 2005, Phys. Rev. D, 71, 103515
Simha V., Weinberg D. H., Davé R., Fardal M., Katz N., Oppenheimer B. D., 2012, MNRAS, 423, 3458
Spergel D. N. et al., 2003, ApJS, 148, 175
Springel V., Yoshida N., White S. D. M., 2001, New Astron., 6, 79
Springel V. et al., 2005, Nat, 435, 629
Tinker J. L. et al., 2012, ApJ, 745, 16
Trujillo-Gomez S., Klypin A., Primack J., Romanowsky A. J., 2011, ApJ, 742, 16
Vale A., Ostriker J. P., 2006, MNRAS, 371, 1173
van den Bosch F. C. et al., 2007, MNRAS, 376, 841
Weinberg D. H., Colombi S., Davé R., Katz N., 2008, ApJ, 678, 6
Zehavi I. et al., 2011, ApJ, 736, 59
Zheng Z., Tinker J. L., Weinberg D. H., Berlind A. A., 2002, ApJ, 575, 617
Zheng Z., Coil A. L., Zehavi I., 2007, ApJ, 667, 760

This paper has been typeset from a TEX/LATEX file prepared by the author.